10. Orthogonalité, projections, bases orthonormales

MTH1008

Sébastien Le Digabel
Polytechnique Montréal

H2024

2024-03-25

v3

Plan

Liens avec le livre et exercices suggérés

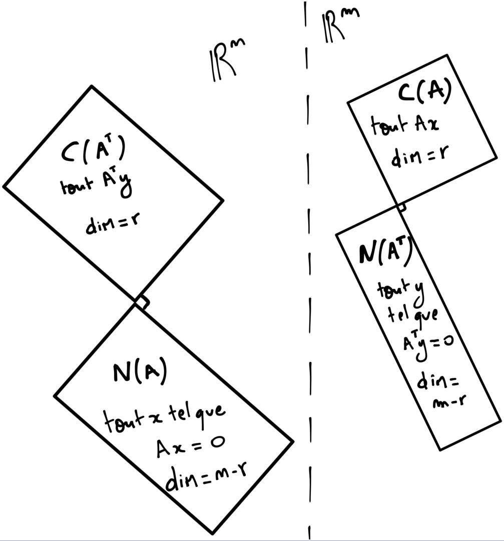
6.1	Orthogonalité	1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 24, 27, 29, 30, 31
6.2	Familles orthogonales	1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 34, 35
6.3	Projections orthogonales	1, 3, 5, 7, 11, 13, 15, 17, 19, 21, 23, 25

Rappels

Si $\mathbf{A}$ est une matrice de taille $m \times n$ et de rang r alors

- $\mathrm{Im}(\mathbf{A}) = C(\mathbf{A})$ est l'image, ou espace des colonnes, engendré par les colonnes de $\mathbf{A}$, et $\dim \mathrm{Im}(\mathbf{A}) = r$. C'est un sous-espace de $\mathbb{R}^m$
- $\mathrm{Ker}(\mathbf{A}) = N(\mathbf{A})$ est le noyau de $\mathbf{A}$, engendré par les solutions spéciales, et $\dim \mathrm{Ker}(\mathbf{A}) = n - r$. C'est un sous-espace de $\mathbb{R}^n$
- $\mathrm{Lgn}(\mathbf{A}) = C(\mathbf{A}^\top)$ est l'espace des lignes, engendré par les lignes de $\mathbf{A}$, et $\dim \mathrm{Lgn}(\mathbf{A}) = r$. C'est un sous-espace de $\mathbb{R}^n$
- $\mathrm{Ker}(\mathbf{A}^\top) = N(\mathbf{A}^\top)$ est le noyau à gauche de $\mathbf{A}$ et $\dim \mathrm{Ker}(\mathbf{A}^\top) = m - r$. C'est un sous-espace de $\mathbb{R}^m$

Portrait global



Vecteurs orthogonaux

- Deux vecteurs $\mathbf{u}$ et $\mathbf{w}$ sont *orthogonaux* si leur produit scalaire est nul
- Dans ce cas, on note $\mathbf{u} \perp \mathbf{w}$, et on a :
  - $\mathbf{u}^\top \mathbf{w} = 0$
  - $\|\mathbf{u}\|^2 + \|\mathbf{w}\|^2 = \|\mathbf{u} + \mathbf{w}\|^2$

Sous-espaces orthogonaux

- Deux sous-espaces U et W (d'un même espace vectoriel) sont *orthogonaux* si $\mathbf{u}^\top \mathbf{w} = 0$ pour chaque vecteur $\mathbf{u} \in U$ et chaque vecteur $\mathbf{w} \in W$
- Autrement dit, tous les vecteurs de U sont orthogonaux à tous les vecteurs de W (et vice versa)
- On note $U \perp W$

Remarques

- Si U et W sont des sous-espaces orthogonaux alors leur seul vecteur commun est le vecteur nul $\mathbf{0}$
- Si U et W sont des sous-espaces de $\mathbb{R}^n$ tels que $\dim U + \dim W > n$, alors il ne peuvent pas être orthogonaux
- Ainsi, par exemple, deux plans de $\mathbb{R}^3$ ne peuvent pas être des sous-espaces orthogonaux

Théorèmes

- Si $\mathbf{A}$ est une matrice de taille $m \times n$ alors le noyau $\mathrm{Ker}(\mathbf{A})$ et l'espace des lignes $\mathrm{Im}(\mathbf{A}^\top)$ sont des sous-espaces vectoriels orthogonaux de $\mathbb{R}^n$:

$$\mathrm{Ker}(\mathbf{A}) \perp \mathrm{Im}(\mathbf{A}^\top)$$

- Si $\mathbf{A}$ est une matrice de taille $m \times n$ alors le noyau à gauche $\mathrm{Ker}(\mathbf{A}^\top)$ et l'image $\mathrm{Im}(\mathbf{A})$ sont des sous-espaces vectoriels orthogonaux de $\mathbb{R}^m$:

$$\mathrm{Ker}(\mathbf{A}^\top) \perp \mathrm{Im}(\mathbf{A})$$

Preuves que $\mathrm{Ker}(\mathbf{A}) \perp \mathrm{Im}(\mathbf{A}^\top)$

- Chaque équation du SÉL $\mathbf{A}\mathbf{x} = \mathbf{0}$ signifie que la ligne correspondante de $\mathbf{A}$ est orthogonale à $\mathbf{x}$
- Donc chaque ligne de $\mathbf{A}$ est perpendiculaire à chaque solution de $\mathbf{A}\mathbf{x} = \mathbf{0}$
- Ceci est vrai aussi pour toutes les combinaisons de lignes

On peut aussi le montrer en considérant $\mathbf{x} \in \mathrm{Ker}(\mathbf{A})$ et $\mathbf{v} \in \mathrm{Im}(\mathbf{A}^\top)$:

- $\mathbf{v}$ peut s'écrire comme combinaison des lignes de $\mathbf{A}$
- Donc on peut trouver un $\mathbf{y} \in \mathbb{R}^m$ tel que $\mathbf{v} = \mathbf{A}^\top\mathbf{y}$
- Ainsi $\mathbf{v}^\top\mathbf{x} = (\mathbf{A}^\top\mathbf{y})^\top\mathbf{x} = \mathbf{y}^\top\mathbf{A}\mathbf{x} = \mathbf{0}$: $\mathbf{x} \perp \mathbf{v}$

Preuves que $\mathrm{Ker}(\mathbf{A}^\top) \perp \mathrm{Im}(\mathbf{A})$

On applique les preuves de $\mathrm{Ker}(\mathbf{A}) \perp \mathrm{Im}(\mathbf{A}^\top)$ à $\mathbf{A}^\top$

Ou bien :

- Chaque équation du SÉL $\mathbf{A}^\top \mathbf{y} = \mathbf{0}$ signifie que la colonne correspondante de $\mathbf{A}$ est orthogonale à $\mathbf{y}$
- Donc chaque colonne de $\mathbf{A}$ est perpendiculaire à chaque solution de $\mathbf{A}^\top \mathbf{y} = \mathbf{0}$
- Ceci est vrai aussi pour toutes les combinaisons de colonnes

Complément orthogonal

Définition

Le *complément orthogonal* (ou l'*orthogonal*) d'un sous-espace vectoriel W est le sous-espace $W^\perp$ contenant **tous** les vecteurs orthogonaux à W

Théorème

Si W est un sous-espace vectoriel de V alors

$$\dim W + \dim W^\perp = \dim V$$

Théorèmes

- Si $\mathbf{A}$ est une matrice de taille $m \times n$ alors le noyau $\mathrm{Ker}(\mathbf{A})$ est le complément orthogonal de l'espace des lignes $\mathrm{Im}(\mathbf{A}^\top)$, dans $\mathbb{R}^n$:

$$\mathrm{Im}(\mathbf{A}^\top) = \mathrm{Ker}(\mathbf{A})^\perp \text{ et } \mathrm{Ker}(\mathbf{A}) = \mathrm{Im}(\mathbf{A}^\top)^\perp$$

- Si $\mathbf{A}$ est une matrice de taille $m \times n$ alors le noyau à gauche $\mathrm{Ker}(\mathbf{A}^\top)$ est le complément orthogonal de l'image $\mathrm{Im}(\mathbf{A})$, dans $\mathbb{R}^m$:

$$\mathrm{Im}(\mathbf{A}) = \mathrm{Ker}(\mathbf{A}^\top)^\perp \text{ et } \mathrm{Ker}(\mathbf{A}^\top) = \mathrm{Im}(\mathbf{A})^\perp$$

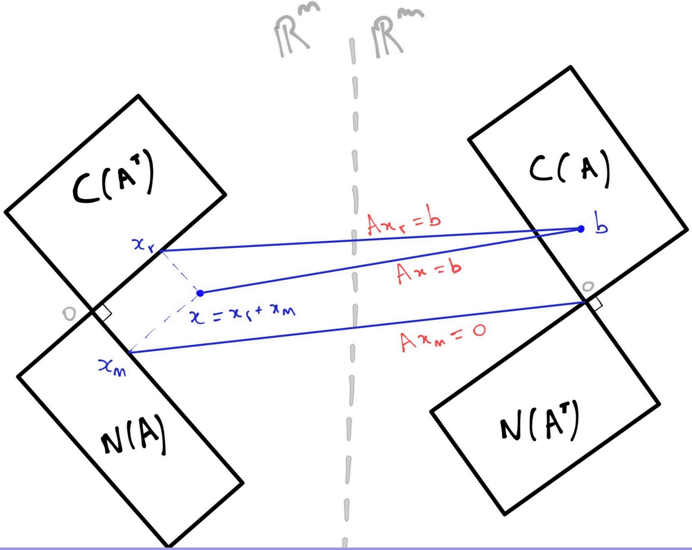
Raison d'être des compléments

Avec une matrice $\mathbf{A} \in \mathbb{R}^{m \times n}$:

- Tout vecteur $\mathbf{x} \in \mathbb{R}^n$ peut se décomposer de façon unique comme
$$\mathbf{x} = \mathbf{x}_r + \mathbf{x}_n$$
avec $\mathbf{x}_r \in \mathrm{Im}(\mathbf{A}^\top)$ et $\mathbf{x}_n \in \mathrm{Ker}(\mathbf{A})$

- Tout vecteur $\mathbf{y} \in \mathbb{R}^m$ peut se décomposer de façon unique comme
$$\mathbf{y} = \mathbf{y}_r + \mathbf{y}_n$$
avec $\mathbf{y}_r \in \mathrm{Im}(\mathbf{A})$ et $\mathbf{y}_n \in \mathrm{Ker}(\mathbf{A}^\top)$

Action de A sur $\mathbf{x}$



Exemple

Avec

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$$

on a

$$\mathbf{x} = \mathbf{x}_r + \mathbf{x}_n$$

avec $\mathbf{x} = (4, 3)$, $\mathbf{x}_r = (2, 4) \in \mathrm{Im}(\mathbf{A}^\top)$, et $\mathbf{x}_n = (2, -1) \in \mathrm{Ker}(\mathbf{A})$

(on calcule $\mathbf{x}_r$ et $\mathbf{x}_n$ par des *projections* de $\mathbf{x}$ sur $\mathrm{Im}(\mathbf{A}^\top)$ et $\mathrm{Ker}(\mathbf{A})$)

Projection sur une droite (1/2)

Soit L le sous-espace vectoriel de $\mathbb{R}^m$ correspondant à la droite engendrée par le vecteur non nul $\mathbf{a} \in \mathbb{R}^m$

- La *projection orthogonale* du vecteur $\mathbf{b} \in \mathbb{R}^m$ sur L est le vecteur $\mathbf{p} \in L$ le plus proche de $\mathbf{b}$
- La projection de $\mathbf{b}$ sur L est $\mathbf{p} = \hat{x}\mathbf{a} = \mathbf{a}\hat{x}$ où $\hat{x} = \frac{\mathbf{a}^\top \mathbf{b}}{\mathbf{a}^\top \mathbf{a}}$
- On peut le voir aussi comme $\mathbf{p} = (\mathbf{u}^\top \mathbf{b})\mathbf{u}$ avec
  - $\mathbf{u} = \frac{\mathbf{a}}{\|\mathbf{a}\|}$ la *normalisation* de $\mathbf{a}$
  - $\mathbf{a} = \|\mathbf{a}\|\mathbf{u}$ et $\|\mathbf{u}\| = \mathbf{u}^\top \mathbf{u} = 1$ ($\mathbf{u}$ est *unitaire*)
- La droite allant de $\mathbf{p}$ à $\mathbf{b}$ est orthogonale à $\mathbf{a}$: $\mathbf{a} \perp \mathbf{e}$ avec $\mathbf{e} = \mathbf{b} - \mathbf{p}$ ($\mathbf{e}$ est appelé l'*erreur*)
- **Exemple :** Avec $\mathbf{a} = (3, 2)$ et $\mathbf{b} = (5, 5)$, on a $\mathbf{p} = \frac{1}{13}(75, 50)$

Projection sur une droite (2/2)

- Matriciellement, le résultat précédent se reformule comme suit :

  La projection de $\mathbf{b} \in \mathbb{R}^m$ sur L est $\mathbf{p} = \mathbf{P}\mathbf{b}$, où $\mathbf{P} \in \mathbb{R}^{m \times m}$ est la *matrice de projection* symétrique suivante :

$$\mathbf{P} = \frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top\mathbf{a}} = \mathbf{u}\mathbf{u}^\top \text{ avec } \mathbf{u} = \frac{\mathbf{a}}{\|\mathbf{a}\|}$$

$$\text{car } \mathbf{p} = \frac{\mathbf{a}^\top\mathbf{b}}{\mathbf{a}^\top\mathbf{a}}\mathbf{a} = \mathbf{a}\frac{\mathbf{a}^\top\mathbf{b}}{\mathbf{a}^\top\mathbf{a}} = \frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top\mathbf{a}}\mathbf{b} = \mathbf{P}\mathbf{b}$$

- **Exemple :**

  Avec $\mathbf{a} = (3, 2)$ et $\mathbf{b} = (5, 5)$, on a $\mathbf{P} = \frac{1}{13}\begin{bmatrix} 9 & 6 \\ 6 & 4 \end{bmatrix}$

Remarques

- $\mathbf{P}$ symétrique, $r(\mathbf{P}) = 1$, $\mathrm{Im}(\mathbf{P}) = \mathrm{Im}([\mathbf{a}]) = L$
- Si $\mathbf{b}$ est orthogonal à $\mathbf{a}$, alors $\mathbf{a}^\top \mathbf{b} = 0$. La projection est $\mathbf{p} = \mathbf{0}$
- Si $\mathbf{b} = \mathbf{a}$, alors $\hat{x} = 1$. La projection de $\mathbf{a}$ sur lui-même donne $\mathbf{a}$: $\mathbf{Pa} = \mathbf{a}$ (mais $\mathbf{P} \neq \mathbf{I}$)
- Si $\mathbf{b} \in L$, alors $\mathbf{p} = \mathbf{b}$ et $\mathbf{e} = \mathbf{0}$
- Projeter une deuxième fois ne change rien :
$$\mathbf{P}^2 = \mathbf{PP} = (\mathbf{uu}^\top)(\mathbf{uu}^\top) = \mathbf{u}(\mathbf{u}^\top\mathbf{u})\mathbf{u}^\top = \mathbf{uu}^\top = \mathbf{P}$$
- Lorsque $\mathbf{P}$ projette sur un sous-espace $(\mathrm{Im}(\mathbf{P}))$, $\mathbf{I} - \mathbf{P}$ projette sur le sous-espace orthogonal $(\mathrm{Ker}(\mathbf{P}))$
- **Exemple :**

  Illustrer avec $\mathbf{a} = (3, 2)$, $\mathbf{b} = (5, 5)$, et $\mathbf{P} = \frac{1}{13}\begin{bmatrix} 9 & 6 \\ 6 & 4 \end{bmatrix}$

Projection dans un sous-espace

Soit $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n \in \mathbb{R}^m$ n vecteurs linéairement indépendants avec $n < m$ et W le sous-espace de $\mathbb{R}^m$ de dimension n engendré par ces vecteurs

- La *projection orthogonale* du vecteur $\mathbf{b} \in \mathbb{R}^m$ sur W est le vecteur $\mathbf{p} \in W$ le plus proche de $\mathbf{b}$
- La projection de $\mathbf{b}$ sur W est $\mathbf{p} = \mathbf{A}\hat{\mathbf{x}} \in \mathbb{R}^m$ où

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \cdots & \mathbf{a}_n \end{bmatrix} \in \mathbb{R}^{m \times n}$$

  et

$$\hat{\mathbf{x}} = \left(\mathbf{A}^\top \mathbf{A}\right)^{-1} \mathbf{A}^\top \mathbf{b} \in \mathbb{R}^n$$

- $\mathbf{p} = \mathbf{P}\mathbf{b}$ où $\mathbf{P}$ est la *matrice de projection*

$$\mathbf{P} = \mathbf{A}\left(\mathbf{A}^\top \mathbf{A}\right)^{-1} \mathbf{A}^\top \in \mathbb{R}^{m \times m}$$

Remarques (1/2)

- Le SÉL $\mathbf{A}\mathbf{x} = \mathbf{b}$ n'a pas toujours de solution $(m > n)$
- $\mathrm{Im}(\mathbf{A}^\top) = \mathbb{R}^n$ et $\mathrm{Ker}(\mathbf{A}) = \{\mathbf{0}\}$
- La matrice $\mathbf{A}^\top\mathbf{A}$ est symétrique et de taille $n \times n$. Elle est inversible car $r(\mathbf{A}) = r(\mathbf{A}^\top\mathbf{A})$ (preuve en exercice) et $r(\mathbf{A}) = n$
- Intuition pour la formule de $\hat{\mathbf{x}}$: la droite allant de $\mathbf{p}$ à $\mathbf{b}$ est orthogonale au sous-espace $W = \mathrm{Im}(\mathbf{A})$:

$$\mathbf{A}^\top(\mathbf{b} - \mathbf{p}) = \mathbf{A}^\top\mathbf{e} = \mathbf{0} = \mathbf{A}^\top(\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}) \Rightarrow \mathbf{A}^\top\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^\top\mathbf{b}$$

$$\Rightarrow \hat{\mathbf{x}} = (\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top\mathbf{b}$$

- Pour trouver la projection $\mathbf{p}$, il faut résoudre le système $\mathbf{A}^\top\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}^\top\mathbf{b}$, qui possède une solution unique

Remarques (2/2)

- Les matrices de projection $\mathbf{P} \in \mathbb{R}^{m\times m}$ vérifient :
  $\mathbf{P}^2 = \mathbf{P}$, $\mathbf{P}^\top = \mathbf{P}$, $r(\mathbf{P}) = n < m$ ($\mathbf{P}$ est singulière)
- $W = \mathrm{Im}(\mathbf{A}) = \mathrm{Im}(\mathbf{P}) = \mathrm{Im}(\mathbf{P}^\top)$: La projection sur W peut être vue comme la multiplication des vecteurs de $\mathbb{R}^n$ par $\mathbf{A}$ ($\mathbf{Ax} \in W$ pour tout $\mathbf{x} \in \mathbb{R}^n$) ou bien comme la multiplication des vecteurs de $\mathbb{R}^m$ par $\mathbf{P}$ ($\mathbf{Pb} \in W$ pour tout $\mathbf{b} \in \mathbb{R}^m$)
- $\mathrm{Ker}(\mathbf{P}) = \mathrm{Ker}(\mathbf{P}^\top) = \mathrm{Ker}(\mathbf{A}^\top)$
- $\mathbf{e} = \mathbf{b} - \mathbf{p}$ est orthogonal aux vecteurs de $W = \mathrm{Im}(\mathbf{A})$. Il est donc dans les noyaux à gauche de $\mathbf{A}$ et $\mathbf{P}$: $\mathbf{A}^\top\mathbf{e} = \mathbf{0}$ et $\mathbf{P}^\top\mathbf{e} = \mathbf{0}$
- Décomposition de $\mathbb{R}^m$: $\mathbf{b} = \underbrace{\mathbf{p}}_{\in \mathrm{Im}(\mathbf{A})} + \underbrace{\mathbf{e}}_{\in \mathrm{Im}(\mathbf{A})^\perp = \mathrm{Ker}(\mathbf{A}^\top)} \in \mathbb{R}^m$

La projection sur une droite est un cas particulier

Si $n = 1$, on retrouve les formules de la projection sur une droite dans $\mathbb{R}^m$:

$$\begin{aligned}
\mathbf{A} \in \mathbb{R}^{m\times n} &\quad\rightarrow\quad \mathbf{a} \in \mathbb{R}^m \\
\mathbf{p} = \operatorname{proj}_{\operatorname{Im}(\mathbf{A})}(\mathbf{b}) &\quad\rightarrow\quad \mathbf{p} = \operatorname{proj}_{\operatorname{Vect}\{\mathbf{a}\}}(\mathbf{b}) \\
\mathbf{p} = \mathbf{P}\mathbf{b} = \mathbf{A}\hat{\mathbf{x}} \in \mathbb{R}^m &\quad\rightarrow\quad \mathbf{p} = \mathbf{P}\mathbf{b} = \hat{x}\mathbf{a} \in \mathbb{R}^m \\
\hat{\mathbf{x}} = \left(\mathbf{A}^\top\mathbf{A}\right)^{-1}\mathbf{A}^\top\mathbf{b} \in \mathbb{R}^n &\quad\rightarrow\quad \hat{x} = \frac{\mathbf{a}^\top\mathbf{b}}{\mathbf{a}^\top\mathbf{a}} \in \mathbb{R} \\
\mathbf{P} = \mathbf{A}\left(\mathbf{A}^\top\mathbf{A}\right)^{-1}\mathbf{A}^\top \in \mathbb{R}^{m\times m} &\quad\rightarrow\quad \mathbf{P} = \frac{\mathbf{a}\mathbf{a}^\top}{\mathbf{a}^\top\mathbf{a}} \in \mathbb{R}^{m\times m}
\end{aligned}$$

1. Orthogonalité

2. Projections

3. Matrices orthogonales et bases orthonormales

Vecteur orthonormaux

- Les vecteurs $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ de $\mathbb{R}^m$ (avec $n \leq m$) sont *orthonormaux* si

$$\mathbf{q}_i^\top \mathbf{q}_j = \begin{cases} 0 & \text{lorsque } i \neq j \quad \text{(orthogonalité)} \\ 1 & \text{lorsque } i = j \quad \text{(vecteurs \textit{unitaires})} \end{cases}$$

  pour tous les $i, j \in \{1, 2, \ldots, n\}$

- Géométriquement, les vecteurs sont de longueur 1 et perpendiculaires entre eux

- Si deux vecteurs sont orthogonaux, alors ils sont indépendants

Matrices orthogonales (1/2)

- Si les colonnes de la matrice $\mathbf{Q} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix} \in \mathbb{R}^{m\times n}$ sont orthonormales alors $\mathbf{Q}^\top \mathbf{Q} = \mathbf{I} \in \mathbb{R}^{n\times n}$

- La matrice $\mathbf{Q}$ n'est pas nécessairement carrée : $\mathbf{Q}^\top \mathbf{Q} = \mathbf{I} \in \mathbb{R}^{n\times n}$ mais $\mathbf{Q}\mathbf{Q}^\top \in \mathbb{R}^{m\times m} \neq \mathbf{I}$

- Dans le cas où $\mathbf{Q}$ est carrée alors on dit que c'est une *matrice orthogonale* et on a $\mathbf{Q}^\top \mathbf{Q} = \mathbf{Q}\mathbf{Q}^\top = \mathbf{I}$, et donc $\mathbf{Q}^{-1} = \mathbf{Q}^\top$

- Les colonnes d'une matrice orthogonale forment une base de $\mathbb{R}^n$. C'est une *base orthonormale*

Matrices orthogonales (2/2)

- Une matrice orthogonale $\mathbf{Q} \in \mathbb{R}^{n\times n}$ préserve le produit scalaire : Pour tous vecteurs $\mathbf{x}$ et $\mathbf{y}$ de $\mathbb{R}^n$, on a

$$(\mathbf{Q}\mathbf{x})^\top(\mathbf{Q}\mathbf{y}) = \mathbf{x}^\top\mathbf{Q}^\top\mathbf{Q}\mathbf{y} = \mathbf{x}^\top\mathbf{y}$$

- Géométriquement, ceci signifie que les longueurs et les angles sont préservés

- $\|\mathbf{Q}\mathbf{x}\| = \|\mathbf{x}\|$ pour tout $\mathbf{x} \in \mathbb{R}^n$

Matrices orthogonales : Exemples

- La matrice orthogonale la plus évidente est l'**identité** : $\mathbf{Q} = \mathbf{I}$
- Matrices de **rotation** : $\mathbf{Q} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \in \mathbb{R}^{2\times 2}$
- Les matrices de **permutation** sont orthogonales
- Matrices de **réflexion** : Si $\mathbf{u} \in \mathbb{R}^n$ unitaire, on définit $\mathbf{Q}$ par la *transformation de Householder* :

$$\mathbf{Q} = \mathbf{I} - 2\mathbf{u}\mathbf{u}^\top \in \mathbb{R}^{n\times n}$$

$\mathbf{Q}$ est symétrique, orthogonale, et $\mathbf{Q}^2 = \mathbf{I}$. Noter aussi que $\mathbf{u}$ ne reste pas comme une colonne de $\mathbf{Q}$

Exemples avec $\mathbf{u}_1 = (1, 0)$ et $\mathbf{u}_2 = \frac{1}{\sqrt{2}}(1, -1)$

Projection : Cas $n \leq m$

- Si les vecteurs $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n$ de $\mathbb{R}^m$ sont orthonormaux alors la matrice de projection sur le sous-espace de $\mathbb{R}^m$ de dimension n engendré par ces vecteurs $(\mathrm{Im}(\mathbf{Q}))$ se simplifie et devient

$$\mathbf{P} = \mathbf{Q}(\mathbf{Q}^\top \mathbf{Q})^{-1}\mathbf{Q}^\top = \mathbf{Q}\mathbf{Q}^\top \in \mathbb{R}^{m \times m}$$

- $\hat{\mathbf{x}} = (\mathbf{Q}^\top \mathbf{Q})^{-1}\mathbf{Q}^\top \mathbf{b} = \mathbf{Q}^\top \mathbf{b} \in \mathbb{R}^n$ avec $\mathbf{b} \in \mathbb{R}^m$: La solution au sens des moindres carrés devient très facile à calculer : Plus besoin de $(\mathbf{A}^\top \mathbf{A})^{-1}$

- $\mathbf{p} = \mathbf{Q}\hat{\mathbf{x}} = \mathbf{Q}\mathbf{Q}^\top \mathbf{b} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix} \begin{bmatrix} \mathbf{q}_1^\top \mathbf{b} \\ \mathbf{q}_2^\top \mathbf{b} \\ \ldots \\ \mathbf{q}_n^\top \mathbf{b} \end{bmatrix} =$

$$\mathbf{q}_1(\mathbf{q}_1^\top \mathbf{b}) + \mathbf{q}_2(\mathbf{q}_2^\top \mathbf{b}) + \ldots + \mathbf{q}_n(\mathbf{q}_n^\top \mathbf{b}) \in \mathrm{Im}(\mathbf{Q}) \subseteq \mathbb{R}^m$$

Projection : Cas $n = m$

- $\mathrm{Im}(\mathbf{Q}) = \mathbb{R}^n$ et la projection d'un vecteur $\mathbf{b} \in \mathbb{R}^n$ sur $\mathrm{Im}(\mathbf{Q})$ est lui-même : $\mathbf{P} = \mathbf{I}$
- De plus, on a $\mathbf{p} = \mathbf{b} \in \mathbb{R}^n$ et

$$\mathbf{b} = \mathbf{q}_1 \left(\mathbf{q}_1^\top \mathbf{b}\right) + \mathbf{q}_2 \left(\mathbf{q}_2^\top \mathbf{b}\right) + \cdots + \mathbf{q}_n \left(\mathbf{q}_n^\top \mathbf{b}\right)$$

$$= \mathrm{proj}_{\mathrm{Vect}\{\mathbf{q}_1\}}(\mathbf{b}) + \mathrm{proj}_{\mathrm{Vect}\{\mathbf{q}_2\}}(\mathbf{b}) + \cdots + \mathrm{proj}_{\mathrm{Vect}\{\mathbf{q}_n\}}(\mathbf{b})$$

  Ceci est la décomposition de $\mathbf{b}$ dans la base des $\mathbf{q}_j$, $j \in \{1, 2, \ldots, n\}$

- **Exemple :**

$$\begin{aligned} \begin{bmatrix} 2 \\ 3 \end{bmatrix} &= 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ &= \tfrac{5\sqrt{2}}{2} \left(\tfrac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right) + \tfrac{\sqrt{2}}{2} \left(\tfrac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right) \end{aligned}$$

Décompositions dans une base orthogonale (1/2)

- Soit W un sev de $\mathbb{R}^m$ de dimension n, et dont une base est $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$
- Si les vecteurs de $\mathcal{B}$ sont orthogonaux, alors $\mathcal{B}$ est appelée une base orthogonale de W
- $\mathbf{Q} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \cdots \ \mathbf{v}_n] \in \mathbb{R}^{m \times n}$ n'est pas orthogonale mais $\mathbf{Q}^\top \mathbf{Q}$ est diagonale $n \times n$ avec chaque élément diagonal égal à $\|\mathbf{v}_j\|^2$, $j = 1, 2, \ldots, n$
- La projection de $\mathbf{b} \in W = \mathrm{Im}(\mathbf{Q})$ donnera alors

$$\mathbf{b} = \frac{\mathbf{v}_1^\top \mathbf{b}}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\mathbf{v}_2^\top \mathbf{b}}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 + \cdots + \frac{\mathbf{v}_n^\top \mathbf{b}}{\|\mathbf{v}_n\|^2}\mathbf{v}_n$$

$$= \mathrm{proj}_{\mathrm{Vect}\{\mathbf{v}_1\}}(\mathbf{b}) + \mathrm{proj}_{\mathrm{Vect}\{\mathbf{v}_2\}}(\mathbf{b}) + \cdots + \mathrm{proj}_{\mathrm{Vect}\{\mathbf{v}_n\}}(\mathbf{b})$$

(décomposition de $\mathbf{b} \in W$ dans $\mathcal{B}$)

Décompositions dans une base orthogonale (2/2)

Si $\mathbf{b} \in \mathbb{R}^m$ (pas forcément dans W), alors

$$\mathbf{b} = \operatorname{proj}_W(\mathbf{b}) + \operatorname{proj}_{W^\perp}(\mathbf{b})$$

$$= \operatorname{proj}_{\mathrm{Vect}\{\mathbf{v}_1\}}(\mathbf{b}) + \operatorname{proj}_{\mathrm{Vect}\{\mathbf{v}_2\}}(\mathbf{b}) + \cdots + \operatorname{proj}_{\mathrm{Vect}\{\mathbf{v}_n\}}(\mathbf{b}) + \operatorname{proj}_{W^\perp}(\mathbf{b}) \in \mathbb{R}^m$$

À noter cette formule alternative de la projection de $\mathbf{b}$ sur W :

$$\mathbf{p} = \operatorname{proj}_W(\mathbf{b}) = \mathbf{Q}(\mathbf{Q}^\top\mathbf{Q})^{-1}\mathbf{Q}^\top\mathbf{b} = \sum_{j=1}^{n} \operatorname{proj}_{\mathrm{Vect}\{\mathbf{v}_j\}}(\mathbf{b})$$

$$= \sum_{j=1}^{n} \frac{\mathbf{v}_j^\top\mathbf{b}}{\|\mathbf{v}_j\|^2}\mathbf{v}_j$$

Résumé des différentes décompositions

$\mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix} \in \mathbb{R}^{m \times n}$ (colonnes orthogonales) :

$$\mathbf{p} = \mathbf{v}_1 \frac{\mathbf{v}_1^\top \mathbf{b}}{\|\mathbf{v}_1\|^2} + \mathbf{v}_2 \frac{\mathbf{v}_2^\top \mathbf{b}}{\|\mathbf{v}_2\|^2} + \ldots + \mathbf{v}_n \frac{\mathbf{v}_n^\top \mathbf{b}}{\|\mathbf{v}_n\|^2} \in \mathrm{Im}(\mathbf{V}) \subseteq \mathbb{R}^m$$

$\mathbf{Q} = \begin{bmatrix} \mathbf{q}_1 & \mathbf{q}_2 & \cdots & \mathbf{q}_n \end{bmatrix} \in \mathbb{R}^{m \times n}$ (colonnes orthonormales) :

$$\mathbf{p} = \mathbf{q}_1(\mathbf{q}_1^\top \mathbf{b}) + \mathbf{q}_2(\mathbf{q}_2^\top \mathbf{b}) + \ldots + \mathbf{q}_n(\mathbf{q}_n^\top \mathbf{b}) \in \mathrm{Im}(\mathbf{Q}) \subseteq \mathbb{R}^m$$

$$\begin{aligned} \mathbf{b} &= \mathrm{proj}_{\mathrm{Im}(\mathbf{V})}(\mathbf{b}) + \mathrm{proj}_{\mathrm{Ker}(\mathbf{V}^\top)}(\mathbf{b}) \\ &= \mathrm{proj}_{\mathrm{Im}(\mathbf{Q})}(\mathbf{b}) + \mathrm{proj}_{\mathrm{Ker}(\mathbf{Q}^\top)}(\mathbf{b}) = \mathbf{p} + \mathbf{e} \end{aligned}$$

Si $m = n$: $\mathrm{Im}(\mathbf{V}) = \mathrm{Im}(\mathbf{Q}) = \mathbb{R}^n$, $\mathrm{Ker}(\mathbf{V}^\top) = \mathrm{Ker}(\mathbf{Q}^\top) = \{\mathbf{0}\}$, $\mathbf{p} = \mathbf{b}$, $\mathbf{e} = \mathbf{0}$, $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ base orthogonale de $\mathbb{R}^n$, $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_n\}$ base orthonormale de $\mathbb{R}^n$